Точкови и интервални оценки. Проверка на хипотези.

- Точкова оценка на θ: $\hat{\theta}$ - функция от данните/наблюденията (статистика). Добри свойства на $\hat{\theta}$: 1. неизместеност: $E(\hat{\theta}) = \theta$; 2. $\hat{\theta}$ да има малка дисперсия за големи извадки.
  Пример: да се докаже неизместеност на $\bar{X} = \frac{X_1 + \ldots X_n)}{n}$ и $S^2 = \sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}$; $Var\bar{X} = \frac{\sigma^2}{n}$

- Метод на моментите за намиране на точкови оценки: $M_k = \sum_{i=1}^{n} \frac{X_i^k}{n}$ е оценка на $E(X^k)$, т.е. съставят се толкова уравнения, колкото параметъра има за оценяване.
  Пример: Засадени са 5 реда по 20 дървета и на следващата година се преброяват оцелелите дръвчета във всеки ред (18, 17, 15, 19, 20). За оценката на вероятността за оцеляване на едно дръвче получаваме $\hat{p} = \bar{X}/20 = 17.8/20 = 0.89$.

- Метод на максималното правдоподобие: функция на правдоподобие $L(\theta) = \prod_{i=1}^{n} f(x_i)$, $f(x)$ е плътността на X. $\hat{\theta}$ е стойността, която максимизира $L(\theta)$ (или $\ln L(\theta)$).
  Пример: Взети са n проби от водата на една река и са преброени коли бактериите във всяка от тях - поасоново разпределени с параметър k. Да се намери МПО на k.
  Пример: Да се намери МПО за μ и σ^2 на извадка от нормално разпределение.

- Ако $X_1, \ldots X_n$ е случайна извадка от $N(\mu, \sigma^2)$, то $\bar{X} \sim N(\mu, \sigma^2/n)$

- Интервална оценка: $[L_1, L_2]$, такъв, че $P(L_1 \leq \theta \leq L_2) = 1 - \alpha$ се нарича $100(1-\alpha)\%$-ен доверителен интервал за параметъра θ.

- Доверителен интервал за μ при известно σ^2: използваме, че $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1)$. Оттук $\bar{X} \pm z_{\alpha/2}\sigma/\sqrt{n}$ е $100(1-\alpha)\%$-ен доверителен интервал за параметъра μ (когато X е нормално разпределена или за голямо n от ЦГТ).
  Пример: При оценка на действието на даден медикамент за лечение на левкемия е измерено средно време на преживяемост на пациентите след поставяне на диагнозата 13 месеца и дисперсия 9. За

95%-ен доверителен интервал на средното време на преживяемост на пациентите вземали даденото лекарство получаваме $P(\bar(X) - 1.96\sigma/\sqrt{n} \leq \mu \leq \bar(X) + 1.96\sigma/\sqrt{n}) = 0.95$. Ако за $n = 16$ пациента е измерено средно $\bar{x} = 13.88$, то получаваме следния доверителен интервал $[12.41, 15.35]$.

- Интервална оценка на дисперсията: Ако $X_1, \dots X_n$ е случайна извадка от $N(\mu, \sigma^2)$, то $(n-1)S^2/\sigma^2 = \sum\limits_{i=1}^{n}(X_i - \bar{X})^2/\sigma^2 \sim \chi^2_{n-1}$, откъдето $100(1-\alpha)$%-ен доверителен интервал за параметъра σ^2 е $[(n-1)S^2/\chi^2_{\alpha/2}, (n-1)S^2/\chi^2_{1-\alpha/2}]$.
  Пример: Дефинирана е релативна мярка за натоварване на компютърна система (1 за дадена система), според която са направени измервания на кръгъл час на системата в голяма консултантска фирма и те са:

3.4	3.6	4.0	0.4	2.0
3.0	3.1	4.1	1.4	2.5
1.4	2.0	3.1	1.8	1.6
3.5	2.5	1.7	5.1	0.7
4.2	1.5	3.0	3.9	3.0

  За да построим 95%-ен доверителен интервал за дисперсията, ни е необходимо да знаем $s^2 = 1.4075$, $\chi^2_{0.025} = 39.4$, $\chi^2_{0.075} = 12.4$ $(n-1 = 24)$. Оттук $L_1 = 24(1.408)/39.4 = 0.858$, $L_2 = 24(1.408)/12.4 = 2.725$.

- Т-разпределение с n степени на свобода: $T = \frac{Z}{\sqrt{\chi^2_n/n}}$,
  $f(t) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\sqrt{\pi n}}\left(1 + \frac{t^2}{n}\right)^{-(n+1)/2}$.

- Доверителен интервал за μ при неизвестно σ^2: използваме, че $\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim T_{n-1}$. Оттук $\bar{X} \pm t_{\alpha/2}S/\sqrt{n}$ е $100(1-\alpha)$%-ен доверителен интервал за параметъра μ.
  Пример: Направени са следните измервания в мкг/к.м. на серен диоксид в дадена гора, поразена от киселинен дъжд:

52.7	43.9	41.7	71.5	47.6	55.1
62.2	56.5	33.4	61.8	54.3	50.0
45.3	63.4	53.9	65.5	66.6	70.0
52.4	38.6	46.1	44.4	60.7	56.4

Намираме $\bar{x} = 53.92$, $s = 10.07$, $s^2 = 101.480$. За $n-1 = 23$, $t_{0.025} = 2.069$ и доверителният интервал е $53.92 \pm 2.069(10.07)/\sqrt{24}$ или $[49.67, 58.17]$.

- Проверка на хипотези: H_0 - нулева хипотеза, H_1 - алтернативна хипотеза. Отхвърляме или не H_0, в зависимост от стойностите на дадена статистика от параметъра. Вероятността наблюдаваната стойност на тази статистика да попадне в т.нар. *критична област* (в която нулевата хипотеза се отхвърля), въпреки че е изпълнена H_0, се нарича *ниво на доверие* и се означава с α. То се избира предварително. Вероятността стойността на наблюдаваната статистика не попадне в критичната област, въпреки че нулевата хипотеза не е вярна се означава с β.

	H_0 е вярна	H_1 е вярна
отхвърля се H_0	грешка от тип I (с вероятност α)	вярно решение
не се отхвърля H_0	вярно решение	грешка от тип II (с вероятност β)

Пример: В проучване на ефекта от светлоотразяващи табели по пътищата, се е стигнало до предположение, че фаровете на повече от половината от автомобилите не са настроени правилно. За проверка на това твърдение съставяме $H_1 : p > 0.5$, $H_0 : p \leq 0.5 (p = 0.5)$. Проверена е настройката на фаровете на 20 автомобила, като броя на тези с неправилна настойка е X. Нека $\alpha = 0.05$, тогава, ако е изпълнена нулевата хипотеза, X има биномно разпределение с $n = 20$, $p = 0.5$, $EX = np = 10$, т.е. ако е изпълнено, че стойността на X е в определена степен по-голяма от 10, ще отхвърлим нулевата хипотеза. Имаме, че $P(X \geq 14 | p = 0.5) = 1 - P(X \leq 13 | p = 0.5) = 1 - 0.9423 = 0.0577 \approx \alpha$, т.е. можем да отхвърлим H_0, ако стойността на X е в множеството $C = \{14, 15, 16, 17, 18, 19, 20\}$ - критична област, и не можем да я отхвърлим, ако $X \in C' = \{0, 1, \ldots 13\}$. В този случай, ако истинската стойност на параметъра е $p = 0.7$, можем да намерим $\beta = P(X \leq 13 | p = 0.7) = 0.3920$, т.е. нашия тест не прави добро разграничение между $p = 0.5$ и $p = 0.7$. За $p = 0.8$, $\beta = 0.0867$.

- *P-value*: може α да не бъде фиксирано предварително, а да се вземе решение, в зависимост вероятността да бъде наблюдавана стойност на тест статистиката поне колкото е стойността и, ако е вярно H_0 :

$\theta = \theta_0$. Тази вероятност често се нарича p-value (така отхвърляме нулевата хипотеза за малки стойности на p-value).

- Тест за средното (t-тест): двустранен - $H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$, едностранен - $H_0 : \mu = \mu_0$, $H_1 : \mu > (<) \mu_0$
  Пример: Компютърна система се състои от 10 компютъра и един принтер, като средното време за стартиране на системата е 15 минути. Добавени са още 10 компютъра и един принтер и трябва да се провери дали средното време се е променило, т.е. $H_0 : \mu = 15$ срещу $H_1 : \mu \neq 15$. Имаме $\bar{x} = 14.0$, $s = 3$, $n = 30$. $(\bar{x} - 15)/(s\sqrt{30}) = -1.83$, $P(T_{29} \leq -1.699) = 0.05$, $P(T_{29} \leq -2.045) = 0.025$ и понеже наблюдаваната стойност на T-статистиката е между тези две стойности, можем да заключим, че вероятността да се наблюдава стойност поне толкова голяма, колкото наблюдаваната (в положителен или отрицателен смисъл - двустранен тест) е между 0.05 и 0.1, което е достатъчно малко, за да можем да отхвърлим нулевата хипотеза.

- Непараметрични методи:
  - Тест на знаците: Нека $X_1, X_2, \ldots X_n \sim X$. Нека M е медианата и $H_0 : M = M_0$, $H_1 : M < M_0 (> M_0, \neq M_0)$. Разглеждаме $X_i - M_0$ и Q_+ е броя на положителните разлики. Ако е вярна H_0, то Q_+ е биномно разпределена с параметри 1/2 и n и очакването и е $n/2$. Ако $P(Q_+ \leq Q_+^{obs} | n, p = 1/2)$ е твърде малка, отхвърляме H_0.
    Пример: Определен етап от производството на машинна част се изпълнява средно за 55 секунди. Пусната е нова технология, за която се твърди, че намалява това време. Измерени са следните времена за новата технология: 35, 65, 48, 40, 70, 50, 58, 36, 47, 41, 49, 39, 34, 33, 31. $P(Q_+ \leq 3 | n = 15, p = 1/2) = 0.0176$, следователно хипотезата за равенство се отхвърля и можем да твърдим, че $M < 55$.
  - Тест на Уилкоксън: Като горното, но се вземат предвид големините на $|X_i - M_0|$ и им се дава ранг $R_1, \ldots R_n$, като на най-малката разлика се дава най-малкия ранг - 1. След това на ранговете се поставя занк, съвпадащ със знака на съответната разлика. Тогава при изпълнена H_0, статистиките $W_+ = \sum_{pozitive} R_i$ и $|W_-| = \sum_{negative} |R_i|$ ще бъдат приблизително

равни. За $W = min(W_+, W_-)$ има таблици и нулевата хипотеза се отхвърля, когато получената стойност на W е по-малка или равна на съответната критична стойност в таблицата.
Пример: Тества се точката на топене на нов материал за интериор на автомобили, като се счита, че медианата е 120^oC. Получени са следните данни: 115.1, 117.8, 116.5, 121.0, 120.3, 119.0, 119.8, 118.5. Потвърждават ли те хипотезата? $W = 5.5$, от таблицата за $\alpha = 0.05, n = 8$, критичната точка е 6, а за $\alpha = 0.025, n = 8$, критичната точка е 4, т.е. можем да отхвърлим нулевата хипотеза и да приемем, че точката на топене е под 120.